



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/600,798	06/20/2003	Hagai Attias	MS302099.1 / MSFTP435US	9699
27195 7590 07/06/2007 AMIN, TUROCY & CALVIN, LLP 24TH FLOOR, NATIONAL CITY CENTER 1900 EAST NINTH STREET CLEVELAND, OH 44114			EXAMINER SIEDLER, DOROTHY S	
			ART UNIT 2626	PAPER NUMBER
			MAIL DATE 07/06/2007	DELIVERY MODE PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary

Application No.

10/600,798

Applicant(s)

ATTIAS ET AL.

Examiner

Dorothy Sarah Siedler

Art Unit

2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 27 April 2007.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-21 is/are pending in the application.
- 4a) Of the above claim(s) 7 is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-6 and 8-21 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 20 June 2003 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
 - ☐ Certified copies of the priority documents have been received in Application No. _____.
 - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☐ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date _____

- 4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____
- 5) ☐ Notice of Informal Patent Application
- 6) ☐ Other: _____

DETAILED ACTION

This office action is in response to the amendment filed April 27, 2007. Claims 1-21 are pending and are considered below.

Response to Amendment

1- The applicant has successfully amended the specification, and as such the objection is withdrawn.

The applicant has successfully amended claims 4,9,10,16 and 18, and as such

The claim objections are withdrawn.

The applicant has successfully overcome the 35 U.S.C. §112 rejection of claim 10, therefore the rejection withdrawn.

Response to Arguments

Applicant's arguments filed April 27th, 2007 have been fully considered but they are not persuasive.

2- The applicant asserts that claims 1,12,14 and 17 "are directed to processes for modeling speech from acoustic data" and that the claims, "recite subject matter that can be applied in a practical application to produce a useful, concrete, and tangible result" (Remarks page 18). The examiner acknowledges that the claimed limitations are drawn to practical application, however claims 1,12,14 and 17 do not recite a useful, concrete, and tangible result or perform a physical transformation. Therefore the rejection of claims 1-19 under 35 U.S.C. §101 is maintained.

The applicant has successfully amended claim 20, and as such the 35 U.S.C. 101 rejection is withdrawn.

In response to applicant's argument that there is no suggestion to combine the references, the examiner recognizes that obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art. See *In re Fine*, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988) and *In re Jones*, 958 F.2d 347, 21 USPQ2d 1941 (Fed. Cir. 1992). In this case, applicant asserts, "There is clearly no explicit suggestion or motivation to use the segmental switching state space model taught by Ghahramani *et al* in a speech processing application". The examiner respectfully disagrees. **Ghahramani** discloses that the switching state space model can be used in a wide range of disciplines, including signal processing (page 1, Introduction), and cites various references (References) for examples. One such reference, **Digalakis** *et al*, uses a segment model for improved modeling of the dynamics of a speech recognition system as an improvement to traditional Hidden Markov Models. In addition, the speech recognition discipline is a subset of signal processing, therefore **Ghahramani** suggests that these models can be implemented as speech recognition models.

Claim Rejections - 35 USC § 101

35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

3- Claims 1-19 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter.

Claims 1-19 define non-statutory processes because they merely manipulate an abstract idea (mathematical algorithm) without a claimed limitation to a physical transformation or a useful, concrete and tangible result. The disclosed invention has a practical application, i.e. speech recognition. However, the claimed process, a series of steps to be performed on a computer, simply manipulates a mathematical algorithm without a claimed limitation to a useful, concrete or tangible result, or a claimed limitation to a physical transformation. In addition, claims 1-19 as a whole fail to show the transformation or the reduction of subject matter to a different state or thing.

Claims 1-19 as a whole merely manipulate an abstract idea (mathematical algorithm), and are therefore non-statutory.

The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

Claim Rejections - 35 USC § 103

Claims 1-5, 7-13, and 19-21 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Hogden** in view of **Ghahramani** ("Variational Learning for Switching State-Space Models" Ghahramani et al., Neural Computation 2000).

4- As per claims 1, 20 and 21, **Hogden** discloses a system and computer readable medium (claim 1) that facilitates modeling unobserved speech dynamics comprising:

- an input component that receives acoustic data (column 10 line 49);
- a model component that employs the acoustic data to characterize speech, the model component comprising model parameters (*pseudo-articulator positions*) that form a mapping relationship from unobserved speech dynamics (*pseudo-articulator positions*) to observed speech acoustics, the model parameters are employed to decode an underlying unobserved phone sequence of speech based, at least in part, upon a variational learning technique (column 5 lines 10-25 and column 8 lines 5-9, (*training, i.e. adjustment of PDF parameters*)).

However, **Hogden** does not disclose wherein the model component is based, at least in part, upon a hidden dynamic model in the form of a segmental switching state space model. **Ghahramani** discloses a probabilistic time-series model in the form of a segmental switching state space model (page 7 section 3: The Generative Model). In addition, **Ghahramani** discloses that the switching state space model can be used in a

wide range of disciplines, including signal processing. The speech recognition discipline is a subset of signal processing therefore **Ghahramani** suggests that these models can be implemented as speech recognition models.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use a model in the form of a segmental state space model in **Hogden**, since it can accurately represent dynamic phenomena, characterized by a combination of discrete and continuous dynamics, as indicated in **Ghahramani** (introduction), such as speech.

5- As per claim 12, **Hogden** discloses a method that facilitates modeling speech dynamics comprising:

- decoding an unobserved phone sequence of speech from acoustic data based, at least in part, upon a speech model, the hidden model comprising at least two sets of parameters, a first set of model parameters describing unobserved speech dynamics (*pseudo-articulator positions*) and a second set of model parameters describing a relationship between an unobserved speech dynamic vector and an observed acoustic feature vector (column 5 lines 10-33, *the continuity map provides a mapping between a variable and a map position, i.e. a sound type and it's articulation*);
- calculating a posterior distribution based on at least the first set of model parameters (column 5 lines 10-33, *PDF*); and,

- modifying at least one of the model parameters based, at least in part, upon the calculated posterior distribution (column 8 lines 5-9, *adjust parameters of the PDF*).

Hogden does not explicitly disclose a hidden model based upon a hidden dynamic model in the form of a segmental switching state space model and calculating a posterior distribution based on the second set of model parameters. However, **Hogden** does disclose that previous attempts to use articulation information to improve speech recognition were based on systems trained with training data consisting of measurements of both articulation data and speech sounds (column 3 lines 8-17). Therefore the examiner argues that it is old and well known to determine the posterior distribution based on the second set of parameters. In addition, **Ghahramani** discloses a probabilistic time-series model in the form of a segmental switching state space model (page 7 section 3: The Generative Model). In addition, **Ghahramani** discloses that the switching state space model can be used in a wide range of disciplines, including signal processing. The speech recognition discipline is a subset of signal processing therefore **Ghahramani** suggests that these models can be implemented as speech recognition models.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to calculate the posterior probability based on the second model parameters in **Hogden**, since speech recognition systems perform more accurately when provided with information about articulator positions and speech sounds, as indicted in **Hogden** (column 2 lines 28-35).

It would also have been obvious to one of ordinary skill in the art at the time of the invention to use a model in the form of a segmental state space model in **Hogden**, since it can accurately represent dynamic phenomena, characterized by a combination of discrete and continuous dynamics, as indicated in **Ghahramani** (introduction), such as speech.

6- As per claim 19, **Hogden** discloses a data packet transmitted between two or more computer components that facilitates modeling of speech dynamics, the data packet comprising: a data structure associated with one or more recovered speech parameters, a speech model based upon acoustic data and model parameters, and the model parameters including at least one articulation parameter and at least one duration parameter (column 20 lines 54-63 and column 6 lines 53-61, *speech is encoded as a pseudo-articulator path, or position, the path including articulator position during a particular time (articulation and duration)*). However, **Hogden** does not disclose a segmental switching state space speech model. **Ghahramani** discloses a probabilistic time-series model in the form of a segmental switching state space model (page 7 section 3: The Generative Model). In addition, **Ghahramani** discloses that the switching state space model can be used in a wide range of disciplines, including signal processing. The speech recognition discipline is a subset of signal processing therefore **Ghahramani** suggests that these models can be implemented as speech recognition models.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use a model in the form of a segmental state space model in **Hogden**, since it can accurately represent dynamic phenomena, characterized by a combination of discrete and continuous dynamics, as indicated in **Ghahramani** (introduction), such as speech.

7- As per claim 2, **Hogden** in view of **Ghahramani** disclose the system of claim 1, and **Hogden** further discloses a modification of at least one of the model parameters being based upon a variational Expectation Maximization algorithm having an E-step and M-step (column 8 line 45 –51, *a path that maximizes the conditional probability data is determined*).

8- As per claim 3, **Hogden** in view of **Ghahramani** disclose the system of claim 2, and **Hogden** further discloses a modification of at least one of the model parameters being based, at least in part, upon a mixture of Gaussian (MOG) posteriors based on a variational technique (column 9 lines 14-17).

9- As per claim 4, **Hogden** in view of **Ghahramani** disclose the system of claim 3, however **Hogden** does not disclose the model component being based, at least in part, upon: the recited equation. **Ghahramani** discloses the use of a probability approximation equation comprising a product or probabilities (page 7, Section 3: The

Art Unit: 2626

Generative Model, equation 2). The equation of the instant application is the standard joint probability equation, modified for independent frames to produce a product of probabilities.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the equation, as noted previously, in **Hogden**, since it is an established formula used within the statistics discipline, therefore enabling the use of readily available software products or algorithms designed for its use.

10- As per claim 5, **Hogden** in view of **Ghahramani** disclose the system of claim 2, and **Hogden** further discloses a modification of at least one of the model parameters being based, at least in part, upon a mixture of hidden Markov model (HMM) posteriors based on a variational technique (column 6 lines 24-46).

11- As per claim 8, **Hogden** in view of **Ghahramani** disclose the system of claim 1, and **Hogden** further discloses the model component being based, at least in part, upon a hidden dynamic model (Abstract, *probabilistic mapping between speech sounds and articulator positions*)

12- As per claim 9 and 10, **Hogden** in view of **Ghahramani** disclose the system of claim 7, and **Ghahramani** further disclose the model component employing, at least in

part, the state equation: the recited equation, and probability distributions the recited equation (page 2, Section 2.1 State-space model, equations (5) and (3) and equation (1)).

Therefore it would have been obvious to one of ordinary skill in the art to use the equation, as noted previously, in **Hogden**, since it would accurately model the input and output behavior of a system, i.e. the conditional probability of an output given a specific input, as indicated in **Ghahramani** (page 2 section 2.1 State-space models).

13- As per claim 11, **Hogden** in view of **Ghahramani** disclose a speech recognition system employing the system of claim 1 **Hogden** (column 1 lines 24-26).

14- As per claims 13, **Hogden** in view of **Ghahramani** disclose the method of claim 12 further comprising receiving the acoustic data **Hogden** (column 10 line 49).

Claims 6 and 14-18 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Hogden** in view of **Ghahramani** as applied to claim 1, and further in view of **McDonough** (5,652,748).

15- As per claim 6, **Hogden** in view of **Ghahramani** disclose discloses the system of claim 1, and **Hogden** further discloses the model component selecting an approximate posterior distribution relating to the acoustic data (column 5 lines 10-33). However,

neither **Hogden** nor **Ghahramani** disclose optimizing a posterior distribution by minimizing a Kullback-Leibler (KB) distance thereof to an exact posterior distribution.

McDonough discloses the use of the Kullback-Leibler distance to determine a likely sequence of words or phrases in a speech recognition system (column 11 lines 40-63).

In addition, **McDonough** discloses that the Kullback-Leibler distance is known in the art, and one of many types of probability models used, any of which would produce an accurate and useful result.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to modify at least one of the model parameters based, at least in part, upon the calculated approximated posterior distribution and minimization of a Kullback-Leibler distance of the approximation from an exact posterior distribution in **Hogden** and **Ghahramani**, since the Kullback-Leibler distance is one of many probability models commonly used, therefore enabling the use a readily available software products or algorithms designed for its use.

16- As per claim 14, **Hogden** discloses a method that facilitates modeling speech dynamics comprising: calculating an approximation of a posterior distribution based on model parameters, the model parameters and the approximation based upon a mixture of Gaussians (column 9 lines 14-17). However, **Hogden** does not disclose recovering speech from acoustic data based, at least in part, upon a speech model in the form of segmental switching state space model and, modifying at least one of the model parameters based, at least in part, upon the calculated approximated posterior

Art Unit: 2626

distribution and minimization of a Kullback-Leibler distance of the approximation from an exact posterior distribution. **Ghahramani** discloses recovering speech from acoustic data based, at least in part, upon a speech model in the form of segmental switching state space model (page 7 section 3: The Generative Model). **Ghahramani** also discloses that the switching state space model can be used in a wide range of disciplines, including signal processing. The speech recognition discipline is a subset of signal processing therefore **Ghahramani** suggests that these models can be implemented as speech recognition models. In addition, **McDonough** discloses modifying at least one of the model parameters based, at least in part, upon the calculated approximated posterior distribution and minimization of a Kullback-Leibler distance of the approximation from an exact posterior distribution (column 11 lines 40-47). **Hogden**, **Ghahramani** and **McDonough** all disclose systems that model observations in relation to states, or hidden states, for the purpose of speech recognition.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to recover speech from acoustic data based, at least in part, upon a speech model in the form of segmental switching state space model and, modifying at least one of the model parameters based, at least in part, upon the calculated approximated posterior distribution and minimization of a Kullback-Leibler distance of the approximation from an exact posterior distribution in **Hogden**, since a segmental switching state-space model can accurately represent dynamic phenomena, characterized by a combination of discrete and continuous dynamics, as indicated in

Ghahramani (introduction), such as speech, and the Kullback-Leibler distance is one of many probability models commonly used, therefore enabling the use of readily available software products or algorithms designed for its use.

17- As per claim 15, **Hogden** in view of **Ghahramani** further in view of **McDonough** disclose the method of claim 14, and **Hogden** further discloses receiving the acoustic data (column 10 line 49).

18- As per claim 16, **Hogden** in view of **Ghahramani** further in view of **McDonough** disclose the method of claim 14, and **Ghahramani** further discloses calculation of the approximation of the posterior distribution being based, at least in part, upon: (see equation claim 16) (page 7, Section 3: The Generative Model). **Ghahramani** discloses the use of a probability approximation equation comprising a product or probabilities (page 7, Section 3: The Generative Model). In addition, the equation of the instant application is the standard joint probability equation, modified for independent frames to produce a product of probabilities. The joint probability equation has been used in the discipline of statistics for many years, and is an established and well known equation.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the equation, as noted previously, in **Hogden**, since it is an established formula used within the statistics discipline which is an effective way to determine the chances of two events occurring at the same time.

19- As per claim 17, **Hogden** discloses a method that facilitates modeling speech dynamics comprising:

- recovering a phone sequence of speech from acoustic data based, at least in part, upon a speech model (column 6 lines 24-46);
- calculating an approximation of a posterior distribution based on model parameters, the model parameters and the approximation based upon a hidden Markov model posteriors (column 6 lines 24-46);

However, **Hogden** does not disclose a speech model in the form of a segmental switching state space model, and modifying at least one of the model parameters based, at least in part, upon the calculated approximated posterior distribution and minimization of a Kullback-Leibler distance of the approximation from an exact posterior distribution. **Ghahramani** discloses a probabilistic time-series model in the form of a segmental switching state space model (page 7 section 3: The Generative Model). In addition, **Ghahramani** discloses that the switching state space model can be used in a wide range of disciplines, including signal processing. The speech recognition discipline is a subset of signal processing therefore **Ghahramani** suggests that these models can be implemented as speech recognition models. **McDonough** discloses the use of the Kullback-Leibler distance is used to determine a likely sequence of words or phrases in a speech recognition system (column 11 lines 40-63). In addition, **McDonough** discloses that the Kullback-Leibler distance is known in the art, and one of many types of probability models used, any of which would produce an accurate and useful result.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use a model in the form of a segmental state space model in **Hogden**, since it can accurately represent dynamic phenomena, characterized by a combination of discrete and continuous dynamics, as indicated in **Ghahramani** (introduction), such as speech.

It would also have been obvious to one of ordinary skill in the art at the time of the invention to modifying at least one of the model parameters based, at least in part, upon the calculated approximated posterior distribution and minimization of a Kullback-Leibler distance of the approximation from an exact posterior distribution in **Hogden**, since the Kullback-Leibler distance is one of many probability models commonly used, therefore enabling the use of readily available software products or algorithms designed for its use.

20- As per claim 18, **Hogden** in view of **Ghahramani** further in view of **McDonough** disclose the method of claim 17, and **Hogden** further discloses calculation of the approximation of the posterior distribution being based, at least in part, upon: the recited equation, where x is a state of the model, s is a phone index, n is a frame number, and, q is a posterior probability approximation. (column 9 lines 27-28, Equation 2, *which for conditional independence among frames, becomes the same function as the equation in the instant application*).

Conclusion

Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).


A shortened statutory period for reply to this final action is set to expire **THREE MONTHS** from the mailing date of this action. In the event a first reply is filed within **TWO MONTHS** of the mailing date of this final action and the advisory action is not mailed until after the end of the **THREE-MONTH** shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than **SIX MONTHS** from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Dorothy Sarah Siedler whose telephone number is 571-270-1067. The examiner can normally be reached on Mon-Thur 9:30am-5:30pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on 571-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

DSS


R. Dorvil
SPE, 266